

# IMPROVING HMM-BASED EXTRACTIVE SUMMARIZATION FOR MULTI-DOMAIN CONTACT CENTER DIALOGUES



Ryuichiro Higashinaka, Yasuhiro Minami, Hitoshi Nishikawa, Kohji Dohsaka, Toyomi Meguro, Satoshi Kobashikawa, Hirokazu Masataki, Osamu Yoshioka, Satoshi Takahashi, Genichiro Kikui

NTT Corporation, Japan

## Overview

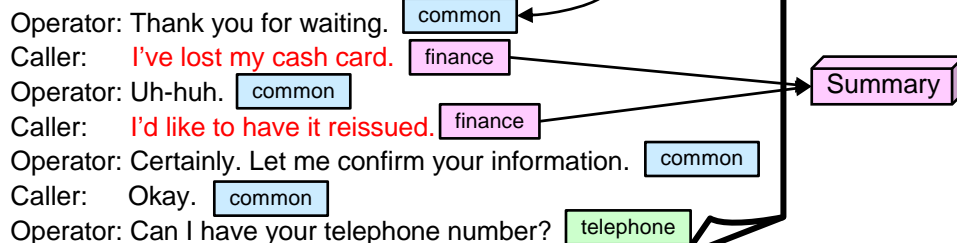
- Objective: **Extractive summarization of multi-domain contact center dialogues** (Domains: finance, PC support, ISP, telephone, mail order, etc.)
- Our previous method using an HMM **could not control compression rates**
- We realized **compression-rate control and better summaries** by a new ILP formulation that uses state posterior probabilities as importance of utterances

## Previous method: HMM-based extractive summarization

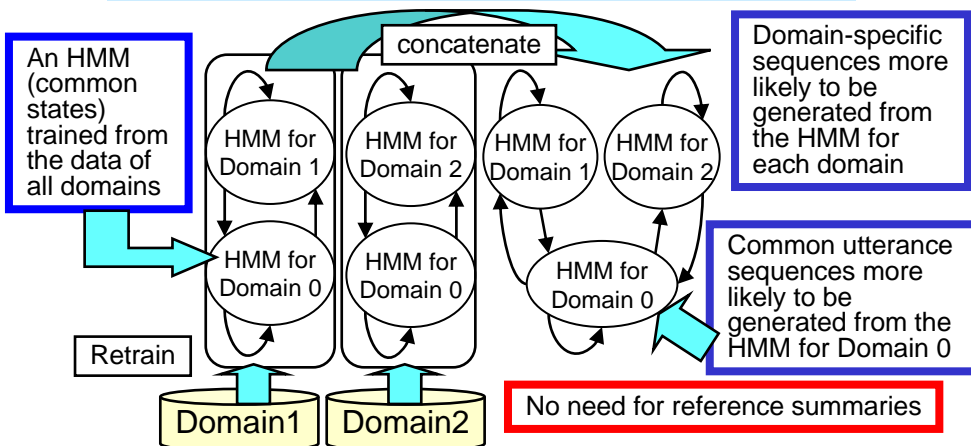
Extract utterances related to a given domain to make a summary

HMM assigns a domain label by Viterbi decoding

Domain: finance



## Our HMM using common states and concatenated training



## Improvements: New ILP formulation

- Maximum coverage of important words under length constraints
- Importance of words is influenced by the importance of utterances

Objective function

Importance of an utterance estimated by the state posterior probability (how likely an utterance relates to a given domain)

$$\operatorname{argmax}_z \sum_i \sum_j m_{ij} w_{ij} z_{ij}$$

$$w_{ij} = \text{weight}(U_i) \cdot \text{weight}(w_j)$$

$W_{ij}$ : weight of  $j$ -th word in  $i$ -th utterance

Importance of a word (eg. Term frequency)

$Z_{ij}$ : whether to include  $j$ -th word in  $i$ -th utterance in a summary

$m_{ij}$ : whether to include  $j$ -th word in  $i$ -th utterance in maximization calculation

## Experiment and Results

Data and References

Training data: 391 dialogues in six domains

Test data: 240 dialogues in six domains (40 dialogues for each domain)

References: 250/500 character extractive summaries made by a human

Evaluation measure: Utterance retrieval accuracy in F-measure

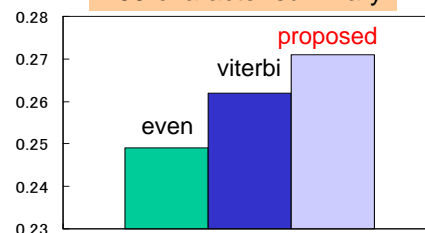
Baselines

even: Use the same value (1.0) for  $\text{weight}(U_i)$

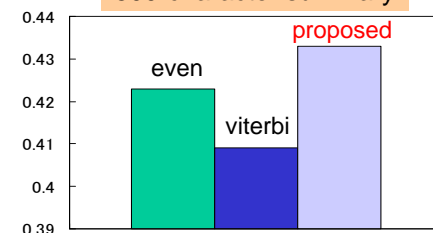
viterbi: Use 1.0 for  $\text{weight}(U_i)$  when the domain label assigned by Viterbi decoding matches the domain of a dialogue; otherwise 0.0

$\text{weight}(W_j)$ : domain-relatedness of a word by the log likelihood ratio

250-character summary



500-character summary



Proposed method made significant improvement for 500-character summaries